



ARTICLE



<https://doi.org/10.1057/s41599-020-0483-7>

OPEN

Skills and motivations underlying children's cumulative cultural learning: case not closed

E. Reindl^{1,2,6}✉, A. L. Gwilliams^{1,3,6}, L. G. Dean², R. L. Kendal^{4,7} & C. Tennie^{5,7}

ABSTRACT The breakthrough study of Dean et al. (Science 335:1114–1118, 2012) claimed that imitation, teaching, and prosociality were crucial for cumulative cultural learning. None of their child participants solved the final stage of their puzzlebox without social support, but it was not directly tested whether the solution was beyond the reach of individual children. We provide this missing asocial control condition, showing that children can reach the final stage of the puzzlebox without social support. We interpret these findings in the light of current understanding of cumulative culture: there are currently conflicting definitions of cumulative culture, which we argue can lead to drastically different interpretations of (these) experimental results. We conclude that the Dean et al. (Science 335:1114–1118, 2012) puzzlebox fulfils a *process-focused* definition, but does not fulfil the (frequently used) *product-focused* definition. Accordingly, the precise role of social support for the apparent taxonomic distribution of cumulative culture and its ontogeny warrants further testing.

¹School of Psychology, University of Birmingham, Birmingham, UK. ²School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, UK. ³School of Life and Health Sciences, Aston University, Birmingham, UK. ⁴Durham Cultural Evolution Research Centre, Department of Anthropology, Durham University, Durham, UK. ⁵Department for Early Prehistory and Quaternary Ecology, University of Tübingen, Tübingen, Germany. ⁶These authors contributed equally: E. Reindl, A. L. Gwilliams. ⁷These authors jointly supervised this work: R. L. Kendal, C. Tennie. ✉email: eva.reindl@live.de

Introduction

A potentially defining feature of humans is the ability to produce *cumulative culture* (CC), a key factor differentiating us from non-humans (Richerson and Boyd, 2005; Tennie et al., 2009; Tomasello et al., 1993; Whiten, 2017, but see Claidière et al., 2014; Jesmer et al., 2018; Sasaki and Biro, 2017; Schofield et al., 2017), and a phenomenon resulting from a nexus of capacities that are believed to be more developed in humans compared to other species, such as language, prosociality or perspective-taking (Tomasello, 2019). Currently, there is no widely agreed-upon definition of CC, a situation which—as we show below—has major implications for both the design and interpretation of experimental findings. Yet, with an increasing number of experimental studies on the social and cognitive processes underlying CC in humans and non-human animals (Caldwell and Millen, 2008; Caldwell et al., 2012; Claidière et al., 2014; Dean et al., 2012; Derex et al., 2013, 2019; Derex and Boyd, 2015; Fay et al., 2019; Jesmer et al., 2018; Mcguigan et al., 2017; Reindl et al., 2017; Sasaki and Biro, 2017; Schofield et al., 2017; Wasieleski, 2014; Zwirner and Thornton, 2015), researchers have highlighted the need for conceptual refinements and clarifications (Caldwell et al., 2016; Charbonneau, 2018; Mesoudi and Thornton, 2018; Miton and Charbonneau, 2018; Reindl et al., 2017; Schofield et al., 2017). One factor that has contributed to misunderstandings and disagreements regarding what constitutes CC, and to what extent it can be found in non-human animals, is the fact that the term “cumulative culture” is used to refer to both cultural *products* (i.e., behaviour or products of behaviour) and *processes* (i.e., cumulative cultural learning, cumulative cultural evolution; Tennie et al., 2018; see Mesoudi and Thornton, 2018, for an overview of current definitions of CC)¹.

For some researchers, CC describes a *process* of a gradual increase in the efficiency and/or complexity of a cultural trait through repeated innovation and transmission events (often over generations; Dean et al., 2014; Mcguigan et al., 2017; Mesoudi and Thornton, 2018; Schofield et al., 2017). Within this framework, the actual level of efficiency/complexity of the final product resulting from such a process is not relevant for determining whether the process is deemed cumulative. Therefore, we will refer to this definition as *process-focused* (Fig. 1a). For example, Schofield et al. (2017) suggested that food-washing behaviours in Japanese macaques has increased in efficiency over a period of 60 years and might thus represent a case of cumulative cultural evolution in non-human primates. The question of whether the most efficient observed technique of food-washing—digging a separate pool of water for rinsing potatoes—could have been invented from scratch without dependency on the older, less efficient techniques—was therefore not decisive for the authors’ conclusion for CC. Another example is the increase in flight distance of paper planes that has been observed in transmission chain experiments with human adults (Caldwell and Millen, 2008, 2009): while the authors acknowledge that the flight distances achieved at the end of the transmission chains could also have been reached by a few individuals without the opportunity of social learning within the experiment, they argue that the observed increase in flight distance, via cycles of learning and innovation, validate the experiment as a laboratory model of cumulative cultural evolution. Such process-focused definitions of CC correspond to the recently defined “core criteria” of CC by Mesoudi and Thornton (2018).

In contrast, other researchers, in addition to describing CC as a process of gradual increase in the efficiency/complexity of a cultural trait, require the efficiency/complexity of the trait (i.e., the *product* of this gradual increase) to go beyond the limits of what any individual of the species could re-innovate from scratch (i.e., on their own; Aplin, 2019; Boyd and Richerson, 1995; Boyd

et al., 2011; Charbonneau, 2015; Henrich and Tennie, 2017; Miton and Charbonneau, 2018; Reindl et al., 2017; Richerson and Boyd, 2005; Tennie et al., 2018, 2009; Tomasello et al., 1993; Vale et al., 2017; Fig. 1b). We refer to this definition as *product-focused*. While for product-focused researchers the criterion of the trait being highly likely to be impossible to innovate by a single individual is a necessary part of the definition of CC, process-focused researchers would see this as an additional feature that only some cumulative cultural products possess (Mesoudi and Thornton, 2018). Note that while we claim that scholars generally fall into one of these categories of defining CC (and some may use both, see e.g., Fay et al., 2019, 2018), we do not suggest that either of these definitions is currently commonly accepted or that one is followed by the majority of researchers in the field.

It has become increasingly clear that CC is not—and should not be regarded as—a unitary phenomenon (Mesoudi and Thornton, 2018). It is likely that with regard to the debate over whether or not CC is unique to humans, researchers may agree that non-human animals possess CC under a process-focused definition, while (according to current evidence) it seems to be restricted to the human species under a product-focused definition (or Mesoudi and Thornton’s (2018) “extended criteria”). Nevertheless, we argue that the lack of clarity (until recently) in using definitions of CC has led researchers to drastically different interpretations of experimental studies aiming to simulate cumulative cultural evolution and learning, resulting in a need to discuss and consolidate the different interpretations. In order to advance our understanding of CC, explicit definitions are required, as they directly influence study design: different definitions will require different methods and criteria for validation. A question for which this problem has become pronounced is whether young children are also capable of cumulative cultural learning or whether this is the preserve of adults (Dean et al., 2012; Mcguigan et al., 2017; Reindl and Tennie, 2018). As we will argue, and in line with the above, the answer depends on which definition is used.

In the search for possible social and cognitive requirements for cumulative cultural learning, researchers have proposed a series of socio-cognitive processes that—on their own or in conjunction with other factors—might be necessary for CC (Biro et al., 2003; Boyd and Richerson, 1985, 1995; Coussi-Korbel and Frigaszy, 1995; Galef, 1992; Giraldeau and Lefebvre, 1987; Hrubesch et al., 2009; Laland, 2004; S. Reader and Laland, 2001; Tennie et al., 2009; Tomasello, 1996, 1999). In 2012, Dean et al. published a paper that was the first to experimentally investigate—within a single study (and across species)—the role of the eight proposed contributing processes at the time: teaching, imitation, prosociality, language, attention to low-ranking innovators, as well as the extent of scrounging, conservatism, and monopolisation of resources by dominant individuals. To examine which of these factors might underlie CC (as a *process*), the authors presented groups of human children (3 to 4 years of age), chimpanzees, and capuchin monkeys with a novel task, which they termed the “cumulative culture puzzlebox” (Dean et al., 2012, p. 1115) and whose three increasingly reward-providing stages had to be opened sequentially. Dean et al. (2012) found that five of the eight groups of children (each comprising 4 to 5 individuals) contained at least two children who reached the third, i.e., final, stage of the puzzlebox. Crucially, while reaching the final stage, and in contrast to the capuchins and chimpanzees who did not reach this stage (apart from one chimpanzee), the children made use of three of the eight investigated socio-cognitive processes: teaching, imitation, and prosociality. In contrast, children who received no such social support from other children in their group (“naturally occurring” asocial controls) did not solve the third stage. This

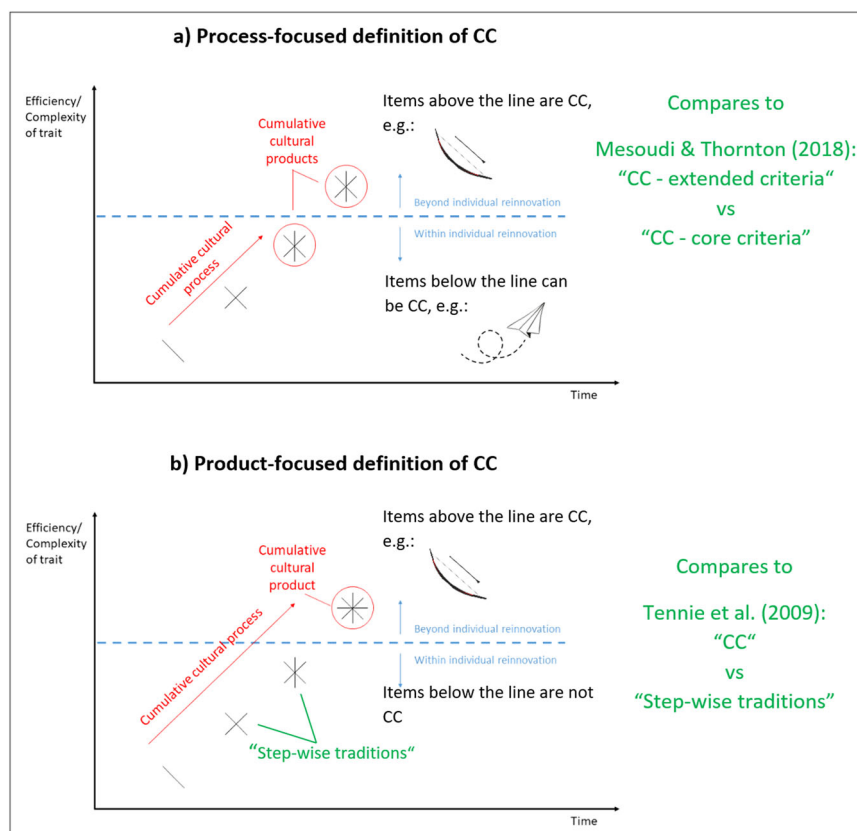


Fig. 1 The two main ways of defining cumulative culture (CC). The line drawings (culminating in "stars") represent cultural traits increasing in efficiency/complexity. **a** The process-focused definition describes CC as a gradual increase in the efficiency/complexity of a cultural trait; the product of such a process is called a cumulative cultural product, regardless of whether the product can be re-innovated by a single "naïve" individual (illustrated by the dashed line). Examples for process-focused CC are the increased flight distance of paper planes over transmission chains of human adults (Caldwell and Millen, 2009) or the improved food-washing behaviours in Japanese macaques (Schofield et al., 2017, both examples lie possibly below the dashed line, i.e., both can potentially be re-innovated by naïve individuals) or a bow and arrow (above the dashed line, probably too complex to be re-innovated from scratch by a single human). The area below the dashed line is equivalent to Mesoudi and Thornton's (2018) *core criteria* CC, the area above is equivalent to their *extended criteria*. *Core criteria* CC is characterised by increases in the learnability of a trait or changes towards a fixed, local optimum (e.g., artificial languages becoming more easily learned (Kirby et al., 2008), pigeon flying routes increasingly approaching optimum (Sasaki and Biro, 2017)), while *extended criteria* CC is open-ended (e.g., many technological products such as ever-improving computers). **b** In the product-focused definition a process/product is labelled as cumulative only when the product of the process is beyond what any "naïve" individual could reinnovate from scratch (i.e., it needs to lie above the dashed line; e.g., bow and arrow). Here, cultural traits that may be individually innovated (such as the paper planes or the food-washing behaviours) are not CC (labelled by some instead as *latent solutions* (Tennie et al., 2009)). For product-focused researchers, CC is inherently open-ended (Tennie et al., 2018, like Mesoudi and Thornton's (2018) *extended criteria* of CC). Here, increases in the learnability of a trait or changes towards local optima of the trait resulting in products that remain within what naïve individuals can re-innovate do not constitute CC, but have been called *step-wise traditions* (Tennie et al., 2009). Note that the term CC only applies to the level of a species or population, but not to the level of an individual. The labels that relate to the level of the individual are those introduced by Lev Vygotsky: the "Zone of Actual Development" (describing what an individual is already capable of doing by themselves) and the "Zone of Proximal Development" (describing what an individual can acquire through social learning). For further discussion on how the Vygotsky's concepts relate to CC, see Reindl et al., (2018).

pattern led the authors to conclude that these three "aspects of human social cognition are directly responsible for the cumulative cultural capability" (ibid., p. 1117).

Dean et al.'s (2012) systematic, multidimensional cross-species approach was a major step forward in the field and their puzzlebox has been accepted as an empirical test for cumulative cultural learning in young children. However, for researchers using a *product-focused* definition of CC, the conclusions of the Dean et al. (2012) study fall short. From the product-focused perspective, in order for the Dean et al. (2012) puzzlebox to be regarded as assessing cumulative cultural learning, its final solution would have to lie beyond the innovative capacities of *individual* children who do not have access to factors such as imitation, teaching, and prosocial acts. In other words, solving all stages of the box has to be shown to require high-fidelity social

learning and hence to represent a "culture-dependent trait" (Reindl et al., 2017) via explicit testing using an appropriately powered asocial control condition (Tennie et al., 2009; see Bordini and Tennie, 2018 for a guideline on how to establish that a given trait is a culture-dependent trait). At the time, due to the "naturally occurring" asocial controls in the Dean et al. (2012) data (see above), an additional asocial control condition of children attempting the task individually was deemed unnecessary. However, while the "naturally occurring" asocial controls in the original study are one form of asocial control condition, for product-focused researchers (McGuigan et al., 2017; Reindl et al., 2017; Vale et al., 2017) they cannot substitute for running a condition with individually-tested participants. First, although the "naturally occurring" asocial control children did not receive teaching, instruction, or prosocial donation of rewards to scaffold

their learning, they may still have observed the task-behaviour of others; thus, we cannot assume that they count as a pure asocial baseline—such a condition is included by design, not deduced a posteriori. Second, the number of these children was, overall, small ($n = 11$)—creating a need for an appropriately powered baseline condition. Third, the lack of social support provided by others (through teaching/prosocial donation of rewards) to these children could have been due to unidentified cues stemming from these “naturally occurring” control children themselves (e.g., being generally uninterested in the task), which itself could be the hidden causal link to their lack of progression in the task. Finally, opportunities to reach the final stage of the task may be higher in an individual baseline condition compared to a group context where individuals necessarily face competition for task access. Overall, there is a need for a dedicated asocial control condition (baseline).

If asocially tested children (baseline) proved able to reach the final stage of the task, this would mean that (1) this task could not be considered a “culture-dependent” trait (*sensu* Reindl et al., 2017) for young children and also that (2) the conclusion of the original study—that imitation, teaching, and prosociality are important for children’s cumulative cultural learning—would have been demonstrated only for a context fitting the process-focused definition of CC. To what extent imitation, teaching, and prosociality are important (or even necessary) for a context that fits the product-focused definition would remain to be empirically demonstrated. For example, Reindl et al. (2017), who endorsed a product-focused view of CC, showed that imitation was not necessary for young children to copy a culture-dependent trait. Here, we provide the asocial control condition (baseline) required to determine whether the full solution of the task also fulfils the product-focused definition of CC. We present children with the Dean et al. (2012) puzzlebox in a truly asocial context, while at the same time aiming to sufficiently motivate children to interact with the box. Thereby, we can determine whether the failure of the “naturally occurring” asocial control children in Dean et al. (2012) to reach the final stage of the puzzlebox was really due to a lack of social support or whether a lack of motivation could have been a potential reason.

Materials and methods

We matched our design to the original Dean et al. (2012) study, but made some changes to increase suitability for testing children individually (see below). Pilot trials with five children, conducted between December 2016 and February 2017, checked for the appropriateness of the trial length and overall practicality of the procedure.

Participants. The final sample consisted of 35 children (age range = 40 to 59 months, $M_{\text{age}} = 51.5$ months, $SD = 5.3$ months) from nine schools and nurseries (see Table S1 for participant characteristics). The sample size of 35 was chosen to match the sample used by Dean et al. (2012). This age group was chosen by the authors of the original study because (1) children of this age have not entered the formal schooling system in the UK yet, thus possessing reduced experience with teaching compared to older children, and (2) because children that age do not possess as much general knowledge yet as older children, thus making the creation of a task that could be a candidate for a culture-dependent product for children more practical. We tested an additional 20 children but had to exclude them due to a script change ($n = 3$), experimenter error ($n = 7$), disruption of the trial by another person ($n = 3$), or because they fell outside the required age range ($n = 7$). Only those children whose parents opted to return completed consent forms participated. All

children were given stickers in reward for participation. Children were allowed to keep all stickers won during their trial; unsuccessful children received six stickers, including two Stage 3 stickers (see below for information on the different kinds of stickers). Ethical approval was received from the STEM Ethical Review Committee at the University of Birmingham.

Materials. We used a warm-up game in order to familiarise children with the testing situation and the experimenter. This was a selection of 15 letter and number dominoes ($34 \times 50\text{mm}$) from the Toys “Я” Us Universe of Imagination Froggy Dominoes Set (Supplementary Fig. S1). Children were encouraged to put the dominoes in numerical order, spell their name, or put the frogs in a queue and knock them over. While Dean et al. (2012) did not include a warm-up session, we deemed this an important part of the current study as children were tested individually and thus might have been more socially inhibited compared to a group testing situation.

For the experiment, children were tested individually in an asocial baseline condition on the same puzzlebox used by Dean et al. (2012). The puzzlebox is symmetrical, consisting of three stages on each side yielding progressively more desirable reward stickers through interaction with the puzzlebox controls (doors, buttons, and dials; Fig. 2). It had two symmetrical, independently-controlled sides and children could solve the stages on either or both of these sides (Supplementary Fig. S2). We used the same stickers as used in the original study (Supplementary Table S2). We used three stopwatches to measure overall trial duration, as well as to independently time the rebait on either side of the box (see below for details on the rebait procedure). Trials were filmed on a Sony HDR-CX330E Handycam mounted on a tripod in a fixed position diagonally behind the participant (Supplementary Fig. S3).

Procedure. Children were tested between March and July 2017 at nurseries and primary schools in West Midlands, England. At each site, a testing location was chosen where other children could not observe the experiment, either in a shielded area of the classroom or another room within the school or nursery. Teachers were asked to prevent participating children from observing others’ trials and from interacting with other participants before it was their turn.

Children were tested individually, in contrast to the group condition used by Dean et al. (2012). Each child participated in one 20 min trial (i.e., the maximum amount of time pilot trials



Fig. 2 Participant interacting with puzzlebox. Four-year-old boy solving the third stage (right side) of the Dean et al. (2012) puzzlebox.

indicated individual children would remain interested in the box. Note that due to practical limitations we gave children less time/opportunity to interact with the puzzlebox than in the original study in which each group received five trials of 30 min each, and that the current study was therefore arguably more conservative). All participants were tested by the same female experimenter (E) using the same procedure. A member of nursery staff was present in the testing room when required by the host institution but was asked not to interact with the child during the trial. Teachers made brief encouraging comments in six trials but these did not corrupt the procedure as the comments resembled the general encouragement comments made by E.

Children were brought individually to the testing area by a member of school staff or E. The child and E sat together at one side of the table and played the warm-up game. During this time, the puzzlebox was on one side of the table, with the front visible to the participant. After a few minutes, the warm-up game was put away and the puzzlebox was moved to the centre of the table (Supplementary Fig. S3). E introduced the puzzlebox and explained that there were three different kinds of stickers in the puzzlebox for them to win. Children were shown the stickers and told that they got better and better (as they increased in size; see Table S2). Before starting the trial, E sat behind the puzzlebox. She then asked the child to try and win some stickers and started the first stopwatch. The full instructions are shown in Table S3.

During the trial, children could manipulate the puzzlebox controls (doors, buttons, and dials) without restriction. This matched the “open” condition in Dean et al. (2012). Participants were asked not to look in the top or back of the puzzlebox. When a child solved a stage and retrieved a sticker, they were requested to put it in a plastic cup. They then continued to try and win more stickers. Details of the trial procedure are shown in Table S4.

Each child participated in a single trial. A trial ended when the main stopwatch reached 20 min ($n = 14$), when a child solved Stage 3 for the second time ($n = 7$), when children stopped engaging with the puzzlebox despite prompting ($n = 13$) or when the child became upset ($n = 1$). Children needed to reach Stage 3 only once to be scored as having solved the full puzzlebox; however, we continued the trial until they reached Stage 3 for the second time or 20 min were over. This was done to demonstrate that the solution was repeatable once learned individually.

Trial durations were measured from the video data. Mean trial duration was 12.94 min (SD = 7.36, range: 1.68–20.78 min). If a child was engaged in an action at the end of a trial, E waited for the child to finish that action ($n = 4$). Experimenter error when setting the stopwatch led to one trial ending 47 s late. See Table S5 for categorised trial durations.

Matching the procedure in Dean et al. (2012), the puzzlebox was rebaited intermittently throughout the trial (2 min after the first task manipulation and 2 min after each rebait). During a rebait the puzzlebox controls (door, button, and dial) were reset to the start position and stickers that had been won were replenished. Rebais on each side of the puzzlebox were timed independently, using separate stopwatches. For each side, a rebait happened 2 min after the first manipulation of any control at this side following the trial start and then 2 min after the first manipulation following each rebait. This matched the Dean et al. (2012) procedure and allowed children to solve stages of the puzzlebox multiple times. See Supplementary Fig. S4 for a detailed timeline of the rebait procedure.

Coded video data indicated that some rebais had not been timed accurately. This occurred for various reasons, such as participant manipulations delaying rebais, errors in identifying first manipulations and multiple demands on E's memory and attention. However, out of the 303 times that the puzzlebox was

rebaited, there were only 6 instances (2%) in which rebait time was more than 50% (60 s) different to the target rebait time of 2 min. It is difficult to hypothesise how rebait time errors could have affected performance. Rebaiting too early may have enabled children to win more stickers over the course of their trial but could also have impeded their progression to higher puzzlebox stages. For example, if they were returned to the start from Stage 2 prematurely, it could take longer to reach Stage 3. Longer rebait times might have had the reverse effect. As the vast majority of rebait times was 2 min, we estimate potential effects of these few longer rebait times to have had a negligible effect. Dean et al. (2012) encountered similar difficulties so it is unlikely that the comparison between the two samples was compromised by these errors.

Statistical analysis. We recorded trial duration, maximum stage reached, and number of stickers won at each stage. Statistical analyses were computed using IBM SPSS Statistics 24.

All data were coded by one of the authors. An independent second rater, naive to the study aims, coded a random 25% selection of the videos (nine videos; 136.17 min). The second coder recorded the number of times participants reached each stage. There was perfect agreement between the two coders ($k = 1$, SE = 0, $p < 0.001$).

We first checked whether there were effects of sex and age on stage reached. We used a two-tailed Mann–Whitney *U*-test to analyse the difference in maximum stage distribution between boys ($n = 14$) and girls ($n = 21$) and found no difference ($U = 151.5$, $z = 0.167$, $p = 0.881$). We analysed the relationship between age and maximum stage reached using a one-tailed Spearman's r and found no relationship ($r_s(33) = 0.62$, $p = 0.362$).

With regard to the comparison of the results of our sample with the results of the original study (both $n = 35$) we used a one-tailed Mann–Whitney *U*-test to analyse the difference in median stage reached between samples; we used a one-tailed test as we hypothesised that if there was a difference between the samples in stage reached, then the children working in groups should show better performance as they could benefit from social learning and prosociality. We used a two-tailed Chi-square test to test for a difference in proportion of children reaching Stage 3 in each study.

We used a two-tailed Chi-square test to compare the proportions of children reaching Stage 1 or higher between both studies.

We also compared our participants with those children in the original study who did not receive any social support: we analysed the difference in median stage reached between all our participants ($n = 35$) and those participants of the original study identified as receiving no social support (“naturally occurring” asocial controls, $n = 11$) using a two-tailed Mann–Whitney *U*-test. We used a two-tailed test as we did not have a directed hypothesis as to why one group should outperform the other (given that all children did the task without social support).

Results

Our key research question was whether *any* child could solve all three stages of the puzzlebox on their own, i.e., without demonstrations or support from others, thus lacking any of the formerly claimed cognitive factors considered necessary for task success (Dean et al., 2012).

Can children solve the puzzlebox on their own? We found that individual children could solve all three stages of the puzzlebox without help: 9 out of the 35 tested children reached the third,

final stage of the puzzlebox at least once, and 7 of these 9 children reached this third stage twice within their trial (the maximum number possible by our design). Our results show clearly that the trio of socio-cognitive processes (teaching, imitation, and prosociality)—while certainly helpful—is *not necessary* for children of this age to solve this particular puzzlebox (which is thus inconsistent with the product-focused definition of CC).

Comparison of the current results with the Dean et al. study

Average performance. We analysed whether children in the original study on average reached higher stages of the puzzlebox than children in the current study. We hypothesised that if there was a difference between the two samples, then the children working in groups (in Dean et al., 2012) should show higher average performance than baseline children (the current study) as they could benefit from teaching, social learning, and prosociality. Indeed, the median stage reached by the individual children tested in the current study (mean rank = 30.16) was significantly lower than that reached by children tested in groups in the original study (one-tailed Mann–Whitney *U*-test, mean rank = 40.84; $U = 425.5$, $z = -2.32$, $p = 0.02$, $\eta^2 = 0.078$; Fig. 3). However, we found no difference in the number of children solving Stage 3 between both studies (two-tailed Chi-square test of independence, $\chi^2 = 2.283$, $p = 0.13$, $\phi = 0.181$). Working in groups therefore increased children's performance overall (consistent with the process-based definition of CC), but did not make them more likely to solve the final stage of the puzzlebox.

Solution of Stage 1 or higher. We also compared both studies with regard to the proportion of children who were able to reach Stage 1 or higher as opposed to those children being “stuck” on Stage 0. We found that the proportion of children who reached Stage 1 or higher (Dean et al.: $n = 27$; current study $n = 16$) compared to those children who remained at Stage 0 (Dean et al.: $n = 8$; current study $n = 19$) was significantly larger in the Dean et al. (2012) than in the current study (two-tailed Chi-square test of independence, $\chi^2 = 7.295$, $p = 0.007$, $\phi = 0.323$), suggesting that social support helped children in being at least minimally successful at the task.

Comparison with “naturally occurring” asocial controls. Finally, we compared performance of children in the current study ($N = 35$) with those children in Dean et al. (2012) who were identified as receiving “no social support” when interacting with the puzzlebox, despite being tested in a group ($n = 11$).

First, we compared whether these two samples differed in the maximum stage reached using a two-tailed Mann–Whitney *U*-test; we found no difference in the distribution of stage reached between the current sample and the Dean et al. (2012) “no social support” children ($U = 231$, $z = 1.103$, $p = 0.333$, $\eta^2 = 0.026$).

Second, we compared whether these two samples differed in the proportion of children reaching Stage 3 (vs. Stage 2 or lower). In Dean et al. (2012), none of the 11 “no social support” children reached Stage 3, whereas nine children in the current study did so. Despite this numerical difference, a two-tailed Fisher's exact test showed that there was no difference between the studies regarding the proportion of these children reaching Stage 3 ($p = 0.064$, $\phi = 0.277$). These two analyses suggest that there was no difference between the performance of our individually tested children and those who received no social support in Dean et al. (2012). Two further analyses regarding the latency to reach Stage 3 can be found in the Supplementary Material.

Discussion

Our main finding was that 9 (26%) of the 35 children we tested in an asocial learning control condition reached the final stage of the Dean et al. (2012) puzzlebox. This was despite them having had less potential time to interact with the puzzlebox (20 min) compared to the original study (5 trials of 30 min) and them being in a more unusual testing situation (1:1 situation with an unfamiliar adult), which might have affected some children's motivation. Our results demonstrate that the trio of socio-cognitive processes (teaching, imitation, and prosociality), while facilitating, was *not necessary* for children's success. Following a product-focused definition (Boyd and Richerson, 1995; Richerson and Boyd, 2005; Tennie et al., 2009; Tomasello et al., 1993), the conclusion is that cumulative cultural evolution was not simulated in Dean et al. (2012) because our results show that the *product* in question (solving all stages of the puzzlebox) was not dependent on any social factors, including social learning, teaching, and prosociality. However, as the groups tested in Dean et al. (2012) showed better performance overall compared to the asocially tested children in

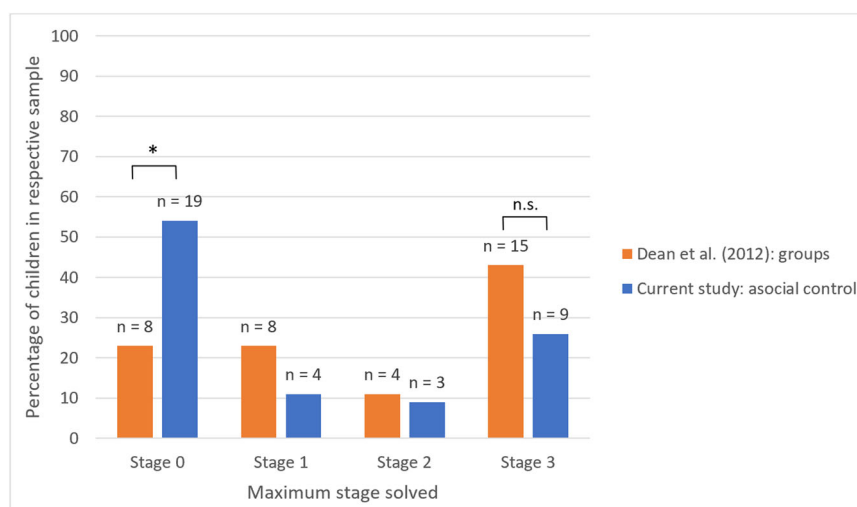


Fig. 3 Percentage of children reaching each stage of the puzzlebox (shown: maximum stage solved). In blue are the results of the current study (children tested individually; one trial of maximum 20 min per child); in orange are the results of the original study by Dean et al. (2012), where children were tested in groups of four or five (5 trials of 30 min per group). Numbers above the bars represent the number of children reaching each maximum stage in the respective study. Total sample size in each study was 35.

the current study, the Dean et al. (2012) puzzlebox does still fulfil the process-focused definition of CC. We conclude that this puzzlebox is invalidated as a proxy for *product*-focused cumulative cultural learning for 3- to 4-year-old children, in part possibly also because the task lacks open-endedness (see also Charbonneau, 2015; Mesoudi and Thornton, 2018). However, this task remains a valid simulation of CC under the *process*-based definition of CC, and may most appropriately be considered as a task illustrating the complexity of a relatively simple ratcheting task for preschool children (i.e., a task whose solution is not culture-dependent (*sensu* Reindl et al., 2017; see also Miton & Charbonneau, 2018, for a discussion on how task complexity affects the ecological validity of and conclusions from laboratory experiments on cumulative cultural evolution).

Children tested in groups showed on average better performance than asocially tested children, and further analysis suggests that social support may have been important in enabling individuals to actually *begin* a seemingly complex task, rather than in ratcheting up their initial solution(s) (Fig. 3). Therefore, the performance difference between the two studies might largely be due to a social facilitation effect (Zajonc, 1965) rather than due to social learning opportunities per se (see also Miton and Charbonneau, 2018): being in a group might have made children more motivated to approach and explore the puzzlebox (although we note that the converse is also true, e.g., some groups invented an alternative game instead of attempting the puzzlebox).

Our findings indicate that the correlation found by Dean et al. (2012) between children's performance and the use of teaching, imitation, and prosociality does not hold in a *product-focused* cumulative cultural context. Indeed, some have suggested that imitation may actually not be necessary for young children to copy a culture-dependent product (Reindl et al., 2017), and the experimental adult literature still debates the roles played by imitation and (imitation-based) teaching in cumulative cultural learning (Caldwell and Millen, 2009; Wasielewski, 2014; Zwirner and Thornton, 2015). Rather, product-focused researchers may now conclude that Dean et al. (2012) demonstrated that some of the socio-cognitive processes claimed to be important or necessary for CC are important for tackling a complex, non-culture-dependent (*sensu* Reindl et al., 2017) ratcheting task. This is perhaps not surprising as there is no reason to assume that those socio-cognitive processes that are argued to support the emergence of culture-dependent traits do not also facilitate the emergence of traits where the problem faced is difficult (though not impossible) to solve asocially (see the predictions of the costly information hypothesis, Kendal et al., 2009).

At first sight it might seem that researchers endorsing the process-focused definition of CC could do without an asocial control condition as we used in the current study: an accumulation in the efficiency and/or complexity of a trait in the experimental condition would seem to suffice to label the observed process CC. However, omitting asocial control conditions can lead to “false positives”, i.e., to the conclusion that one has identified a cumulative cultural process where in fact there was none. For example, in studies exploring the *subtractive ratchet effect* (Tennie et al., 2014; see also McGuigan and Graham, 2009) chains of children that were seeded with a particularly inefficient way of carrying out a certain task (e.g., carrying rice from location A to B, opening a puzzlebox) showed an improvement over generations in their efficiency, which might seem like a cumulative cultural process. However, only a comparison with control groups could show that the behaviour at the end of the chains was not more efficient than what individual children could reinnovate without any demonstration. Thus, this process was a mere recovery from a seeded inefficient technique to baseline performance—a special kind of ratchet effect, but

probably not a process most researchers endorsing either definition of CC would regard as a cumulative cultural process (especially because the improvements were not based on social transmission but on the absence of copying, as children had to refrain from copying the inefficient technique in order to become more efficient). Therefore, regardless of the definition of CC one endorses, any experimental study on cumulative cultural evolution should include an asocial control condition in order to avoid such false positives.

This paper highlights the need for considering different definitions of CC in the field of cultural evolution in general (alongside Mesoudi and Thornton, 2018; Miton and Charbonneau, 2018) and, specifically, regarding the question of whether non-human animals and human children are capable of cumulative cultural learning. Here, we do not endorse one CC definition over the other, as our study was never designed to decide between or test these definitions. Perhaps the most important implication of our results is that the question of whether groups of young children can already produce culture-dependent traits (*sensu* Reindl et al., 2017) themselves and—if so—whether the skills and motivation for imitation, teaching, and prosociality (or other factors) are necessary for children to do so, requires renewed investigation—using tasks whose solutions are beyond what can be achieved individually by the subjects in question. We already know that children readily *copy* culture-dependent traits even when not explicitly asked to do so (Reindl et al., 2017). With regard to *producing* culture-dependent traits, a promising start has already been made (McGuigan et al., 2017; Reindl and Tennie, 2018) but more research is essential for the development of appropriate tasks (Miton and Charbonneau, 2018). Such tasks can be identified by administering candidate tasks to a large number of participants tested in asocial conditions (for extended time periods), which will delineate the asocial learning performance of a given study group (as we did here). If the solution to a task is not spontaneously reinnovated by any of the participants in this baseline (Bandini and Tennie, 2018; Reindl et al., 2017; Tennie et al., 2009), investigations may progress. We further suggest that demonstration of product-focused CC (or extended criteria CC, Mesoudi and Thornton, 2018) should focus upon *open-ended* tasks rather than those with a priorly known optimal solutions. This is especially important as it has been suggested that open-endedness might be a key difference in the cultural evolution of humans and non-human animals (Charbonneau, 2015; Tennie et al., 2018). Yet note that open-endedness is a feature of the process, not of the product, of cumulative cultural evolution. Therefore, it is vital for our understanding of what makes CC unique in humans to not only focus on cumulative cultural products (e.g., whether they increase in efficiency, whether they are beyond what any individual can reinnovate) but to also study the nature of the process itself (see e.g., Charbonneau, 2015).

CC is a hallmark of our species that has allowed us to populate almost all habitats on the planet and even to venture into space (Henrich, 2015). More clarity in the use of definitions and experimental tasks is essential if we want to further our understanding of the roots of human cultural success. One of the major questions to be answered is whether CC also exists in any non-human animals. So far, CC has been viewed as a unitary phenomenon (and a potential rubicon between humans and non-humans; see also Mesoudi and Thornton, 2018); however, it is likely that for this question the answer likewise depends on the definition. We may find that (some) species have CC under a process-based definition (or core criteria CC, Mesoudi and Thornton, 2018), while human culture may be unique in fulfilling also the product-based definition (or extended criteria CC, Mesoudi and Thornton, 2018). In order to study this possibility, explicitly definition-matched tasks that are

appropriate for comparative empirical research are required. Our findings thus pave the way for a more robust investigation of whether the other-regarding socio-cognitive capacities identified by Dean et al. (2012) explain the difference between human and non-human cultural achievements and, if so, when and why these capacities may have evolved and when they appear ontogenetically in humans.

Data availability

The datasets generated during the current study are available in the OSF repository, <https://doi.org/10.17605/OSF.IO/TRG3V>.

Received: 4 December 2019; Accepted: 28 April 2020;

Published online: 28 May 2020

Notes

- 1 Similar issues of whether a phenomenon in question is defined as a process or a product have been encountered, e.g., by innovation researchers (innovation as a product that is novel to the group vs innovation as a process that is novel to the individual; Carr et al., 2016; Reader and Laland, 2003).

References

- Aplin LM (2019) Culture and cultural evolution in birds: a review of the evidence. *Anim Behav* 147:179–187
- Bandini E, Tennie C (2018) Naive, captive long-tailed macaques (*Macaca fascicularis fascicularis*) fail to individually and socially learn pound-hammering, a tool-use behaviour. *R Soc Open Sci* 5:171826
- Biro D, Inoue-Nakamura N, Tonooka R et al. (2003) Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Anim Cogn* 6(4):213–223
- Boyd R, Richerson PJ (1985) Culture and the evolutionary process. Chicago University Press, Chicago
- Boyd R, Richerson PJ (1995) Why culture is common, but cultural evolution is rare. *Proc Br Acad* 88:77–93
- Boyd R, Richerson PJ, Henrich J (2011) The cultural niche: why social learning is essential for human adaptation. *Proc Natl Acad Sci USA* 108:10918–10925
- Caldwell CA, Cornish H, Kandler A (2016) Identifying innovation in laboratory studies of cultural evolution: rates of retention and measures of adaptation. *Philos T R Soc B* 371(1690):20150193
- Caldwell CA, Millen AE (2008) Experimental models for testing hypotheses about cumulative cultural evolution. *Evol Hum Behav* 29(3):165–171
- Caldwell CA, Millen AE (2009) Social learning mechanisms and cumulative cultural evolution: is imitation necessary? *Psychol Sci* 20(12):1478–1483
- Caldwell CA, Schillinger K, Evans CL et al. (2012) End state copying by humans (*Homo sapiens*): implications for a comparative perspective on cumulative culture. *J Comp Psychol* 126(2):161–169
- Carr K, Kendal RL, Flynn EG (2016) Eureka!: what is innovation, how does it develop, and who does it? *Child Dev* 87(5):1505–1519
- Charbonneau M (2015) All innovations are equal, but some more than others: (Re) integrating modification processes to the origins of cumulative culture. *Biol Theory* 10(4):322–335
- Charbonneau M (2018) Understanding cultural fidelity. *Br J Philos Sci* <https://doi.org/10.1093/bjps/axy052>
- Claidière N, Smith K, Kirby S et al. (2014) Cultural evolution of systematically structured behaviour in a non-human primate. *Proc R Soc B* 281(1797):20141541
- Coussi-Korbel S, Frigaszy DM (1995) On the relation between social dynamics and social learning. *Anim Behav* 50:1441–1453
- Dean LG, Kendal RL, Schapiro SJ et al. (2012) Identification of the social and cognitive processes underlying human cumulative culture. *Science* 335:1114–1118
- Dean LG, Vale GL, Laland KN et al. (2014) Human cumulative culture: a comparative perspective. *Biol Rev* 89(2):284–301
- Derex M, Beugin MP, Godelle B et al. (2013) Experimental evidence for the influence of group size on cultural complexity. *Nature* 503(7476):389–391
- Derex M, Bonnefon J-F, Boyd R et al. (2019) Causal understanding is not necessary for the improvement of culturally evolving technology. *Nat Hum Behav* 3:446–452
- Derex M, Boyd R (2015) The foundations of the human cultural niche. *Nat Comm* 6:8398
- Fay N, De Kleine N, Walker B et al. (2019) Increasing population size can inhibit cumulative cultural evolution. *Proc Natl Acad Sci* 116(14):6726–6731
- Fay N, Ellison TM, Tylén K et al. (2018) Applying the cultural ratchet to a social artefact: the cumulative cultural evolution of a language game. *Evol Hum Behav* 39(3):300–309
- Galef BG (1992) The question of animal culture. *Hum Nat* 3(2):157–178
- Giraldeau L-A, Lefebvre L (1987) Scrounging prevents cultural transmission of food-finding behaviour in pigeons. *Anim Behav* 35:387–394
- Henrich J (2015) The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter. Princeton University Press, New Jersey
- Henrich J, Tennie C (2017) Cultural evolution in chimpanzees and humans. In: Muller M, Wrangham R, Pilbeam D (eds) *Chimpanzees and human evolution*. Harvard University Press
- Hrubesch C, Preuschoft S, van Schaik C (2009) Skill mastery inhibits adoption of observed alternative solutions among chimpanzees (*Pan troglodytes*). *Anim Cogn* 12(2):209–216
- Jesmer BR, Merkle JA, Goheen JR et al. (2018) Is ungulate migration culturally transmitted? Evidence of social learning from translocated animals. *Science* 361:1023–1025
- Kendal RL, Kendal JR, Hoppitt W (2009) Identifying social learning in animal populations: a new “option-bias” method. *PLoS ONE* 4(8):e6541. <https://doi.org/10.1371/journal.pone.0006541>
- Kirby S, Cornish H, Smith K (2008) Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proc Natl Acad Sci USA* 105(31):10681–10686
- Laland KN (2004) Social learning strategies. *Anim Learn Behav* 32(1):4–14
- McGuigan N, Burdett E, Burgess V et al. (2017) Innovation and social transmission in experimental micro-societies: exploring the scope of cumulative culture in young children. *Philos T R Soc B* 372:20160425
- McGuigan N, Graham M (2009) Cultural transmission of irrelevant tool actions in diffusion chains of 3- and 5-year-old children. *Eur J Dev Psychol* 7(5):561–577
- Mesoudi A, Thornton A (2018) What is cumulative cultural evolution? *Proc R Soc B* 285:20180712
- Miton H, Charbonneau M (2018) Cumulative culture in the laboratory: methodological and theoretical challenges. *Proc R Soc B* 285(1879):20180677
- Reader S, Laland K (2001) Primate innovation: sex, age and social rank differences. *Int J Primatol* 22(5):787–805
- Reader SM, Laland KN (2003) *Animal innovation*. Oxford University Press, Oxford, UK
- Reindl E, Apperly IA, Beck SR et al. (2017) Young children copy cumulative technological design in the absence of action information. *Sci Rep* 7(1):1788
- Reindl E, Bandini E, Tennie C (2018) The zone of latent solutions and its relation to the classics: Vyotsky and Köhler. In: Di Paolo LD, Di Vincenzo F, De Petrillo F (eds) *Evolution of primate social cognition*. Springer, Cham, pp. 231–248
- Reindl E, Tennie C (2018) Young children fail to generate an additive ratchet effect in an open-ended construction task. *PLoS ONE* 13(6):e0197828
- Richerson PJ, Boyd R (2005) Not by genes alone. How culture transformed human evolution. Chicago University Press, Chicago
- Sasaki T, Biro D (2017) Cumulative culture can emerge from collective intelligence in animal groups. *Nat Comm* 8:15049
- Schofield DP, McGrew WC, Takahashi A et al. (2017) Cumulative culture in nonhumans: overlooked findings from Japanese monkeys? *Primates* 59:0123456789
- Tennie C, Caldwell CA, Dean LG (2018) Cumulative culture. In: Callan H (ed) *International encyclopedia of anthropology*. Wiley-Blackwell
- Tennie C, Call J, Tomasello M (2009) Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos T R Soc B* 364(1528):2405–2415
- Tennie C, Walter V, Gampe A et al. (2014) Limitations to the cultural ratchet effect in young children. *J Exp Child Psychol* 126:152–160
- Tomasello M (1996) The question of chimpanzee culture. In: Wrangham RC, McGrew WC, de Waal FBM, Heltne PG (eds) *Chimpanzee cultures*. Harvard University Press, Cambridge, MA, pp. 301–377
- Tomasello M (1999) *The cultural origins of human cognition*. Harvard University Press, Cambridge, MA
- Tomasello M (2019) *Becoming human. A theory of ontogeny*. Belknap Press, Cambridge, MA
- Tomasello M, Kruger AC, Ratner HH (1993) Cultural learning. *Behav Brain Sci* 16(03):495–552
- Vale GL, Carr K, Dean LG et al. (2017) The cultural capacity of human and nonhuman primates: social learning, innovation, and cumulative culture. In: *Evolution of nervous systems*, 2nd edn. Oxford Academic Press, Oxford, pp. 475–508
- Wasielowski H (2014) Imitation is necessary for cumulative cultural evolution in an unfamiliar, opaque task. *Hum Nat* 25(1):161–179
- Whiten A (2017) Social learning and culture in child and chimpanzee. *Ann Rev Psychol* 68(1):129–154
- Zajonc RB (1965) Social facilitation. *Science* 149(3681):269–274
- Zwirner E, Thornton A (2015) Cognitive requirements of cumulative culture: teaching is useful but not essential. *Sci Rep* 5(1):16781

Acknowledgements

We thank all children and their parents, the nurseries (Beecher Bear, Busy Bees, Ducklings, Halesowen College, Happy Faces, Little Hands, and Perryfields), and primary schools (Glynne, Greenfield, Pedmore, Perryfields, and Rufford) for their help and participation in the studies; Claire Thirkettle, Isabel Williams, Natasha Bargota, and Faryad Yusuf for help with reliability coding and transcription; Kevin Laland, Sally Street, the reading group of the Durham Cultural Evolution Research Centre, Elisa Bandini, and Damien Neadle for very helpful comments on earlier versions of the manuscript. During study design and data collection, the research of CT was supported by a grant from the Economic and Social Research Council (ESRC) (ES/K008625/1). At the time of writing, CT was supported by a grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 714658; STONECULT project). During his work on the Dean et al. (2012) study, LGD was supported by the CULTAPTATION project (European Commission contract FP6-2004-NESTPATH-043434) and a European Research Council Advanced Grant (EVO-CULTURE, 232823) awarded to Kevin Laland. The funding sources had no involvement in study design, collection, analysis, interpretation of the data, writing the paper, nor in the decision to submit the article for publication.

Author contributions

All authors designed the study, ALG collected and analysed the data, all authors discussed the results, ALG, ER, RLK, CT wrote the manuscript, all authors commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1057/s41599-020-0483-7>.

Correspondence and requests for materials should be addressed to E.R.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020